Computational and Structural Biotechnology Journal 20 (2022) 3955–3962

COMPUTATIONAL ANDSTRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Mapping nucleosome and chromatin architectures: A survey of computational methods

Kun Fang [a], Junbai Wang [b], Lu Liu [c], Victor X. Jin [a,*]

[a] Institute for Health and Equity, MCW Cancer Center, and Mellowes Center for Genome Science and Precision Medicine, Medical College of Wisconsin, Milwaukee, WI, USA
[b] Department of Clinical Molecular Biology, University of Oslo and Akershus University Hospital, Lørenskog, Norway
[c] Department of Computer Science, North Dakota State University, Fargo, ND, USA

ARTICLE INFO

Article history:
Received 8 March 2022
Received in revised form 22 July 2022
Accepted 22 July 2022
Available online 26 July 2022

Keywords:
Computational methods
Nucleosome
Chromatin architectures

ABSTRACT

With ever-growing genomic sequencing data, the data variabilities and the underlying biases of the sequencing technologies pose significant computational challenges ranging from the need for accurately detecting the nucleosome positioning or chromatin interaction to the need for developing normalization methods to eliminate systematic biases. This review mainly surveys the computational methods for mapping the higher-resolution nucleosome and higher-order chromatin architectures. While a detailed discussion of the underlying algorithms is beyond the scope of our survey, we have discussed the methods and tools that can detect the nucleosomes in the genome, then demonstrated the computational methods for identifying 3D chromatin domains and interactions. We further illustrated computational approaches for integrating multi-omics data with Hi-C data and the advance of single-cell (sc)Hi-C data analysis. Our survey provides a comprehensive and valuable resource for biomedical scientists interested in studying nucleosome organization and chromatin structures as well as for computational scientists who are interested in improving upon them.

Contents

Abbreviations: 3D, three dimensional; ATAC-seq, assay for transposase-accessible chromatin using sequencing; ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; ChIP-seq, chromatin immunoprecipitation sequencing; CTCF, CCCTC-binding factor; ICE, iterative correction; Mb, megabases; Micro-C, micrococcal nuclease applied chromosome conformation capture; MNase, micrococcal nuclease; MPE-seq, methidiumpropyl-EDTA sequencing; PCA, principle component analysis; PF, pioneer factor; TAD, topologically associating domain; TCC, tethered chromosome conformation capture; TF, transcription factor.

* Corresponding author.
E-mail address: vjin@mcw.edu (V.X. Jin).

https://doi.org/10.1016/j.csbj.2022.07.037

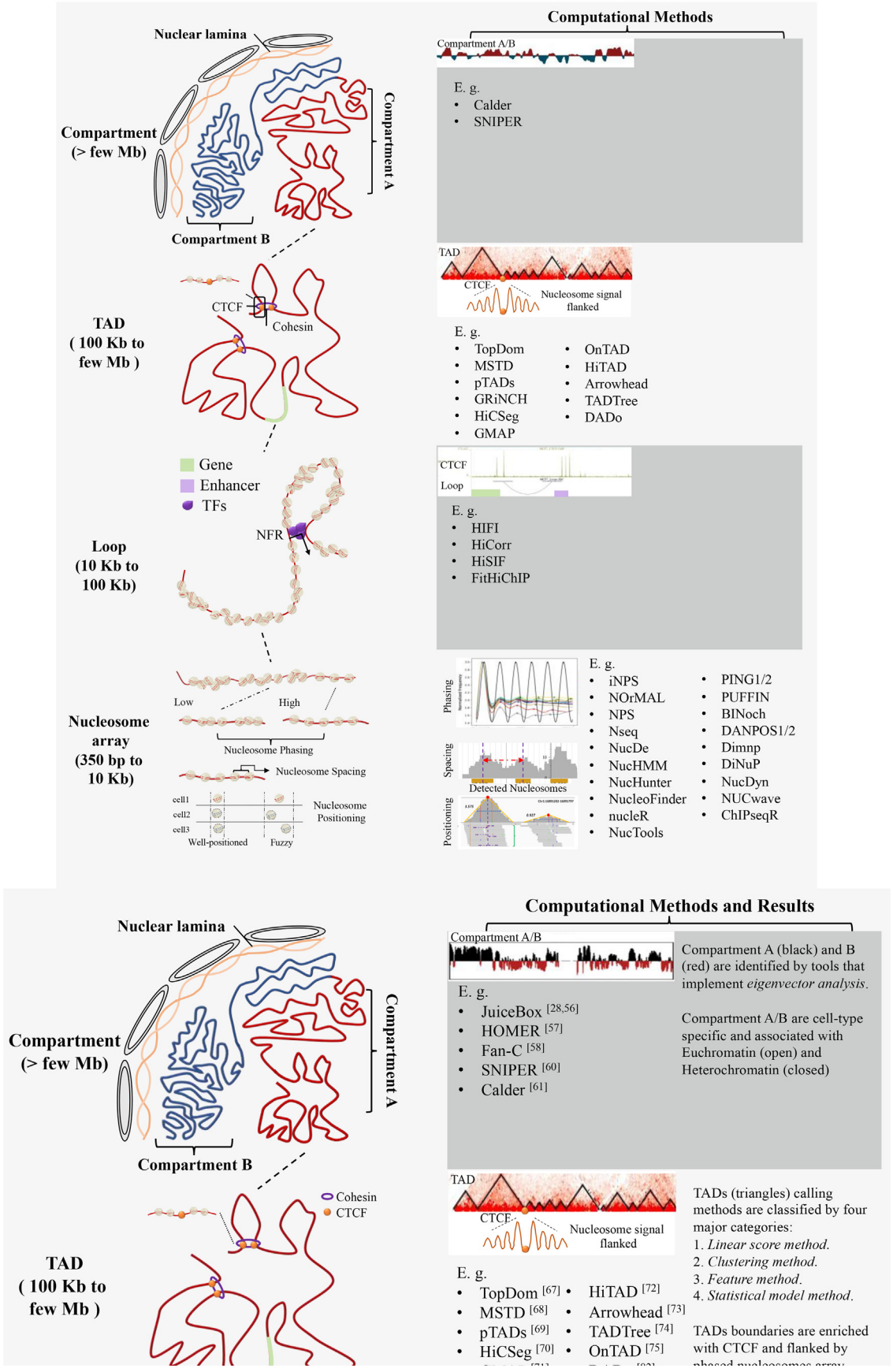

Fig. 1. Multiple scales of DNA folding in the nucleus (cartoons in the left panels), from the small scale organization of nucleosome array to the median scaling structure of loop, and the large scale structure of TAD as well as compartment A/B.

1. Introduction

A nucleosome is the unitary structure of the chromatin fiber, composed of two copies of each of four histone proteins, H2A, H2B, H3, and H4, and a ~ 147 bp DNA wrapped around the histone octamer [1]. Nucleosomes along with other proteins or complexes such as the structural insulator CCCTC-binding factor (CTCF) [2,3] and long-range interaction mediator YY1 [4] facilitate the assembly of three-dimensional (3D) chromatin structure. 3D chromatin structures, ranging from megabases (Mb)-scaled chromatin compartment, kilobases (Kb)-Mb-scaled topologically associating domain (TAD), to Kb-scaled chromatin loop, play critical roles in gene regulation [5], DNA replication [6], cell development and differentiation [7]. Recent advance in sequencing technologies allows us to map the nucleosome organization and 3D chromatin structure (Fig. 1). MNase-seq, a profiling method for mapping nucleosome landscape, is the most prevalent technique to study nucleosome organization. In addition, several other experimental methods were developed to tackle the nucleosome arrangement in the nucleus, including Array-seq [8], Methylation foot-printing [9], Chemical mapping [10], ATAC-seq [11] and MPE-seq [12]. These alternative approaches adapted the idea of cleaving DNA between the nucleosomes but provided extra information. For example, chemical mapping cleaves the DNA by introducing cysteines in histones H3/H4 and can detect nucleosome preferred positions. MPE-seq uses MPE-Fe(II) to digest the DNA and can find the presence of core histones. Chromosome conformation capture (3C) [13] and its derivatives, including but not limit to circular 3C (4C) [14], 3C carbon copy (5C) [15], Capture-C [16], Hi-C [17], TCC [18], and Micro-C [19], have been widely used to study chromatin structures.

With ever-growing genomic sequencing data, the data variabilities and the underlying biases barriered researchers from acquiring biological information accurately. For instance, MNase cleaves DNA about 30 times slower upstream of an G/C than 5′ of a A/T, which causes sequence bias for MNase-seq [20–22]. And Yaffe et al reported systematic bias in Hi-C dataset, including different length of restriction fragments caused by ligation efficiencies, different sequencing amplifications affected by GC contents, and differential mappability of sequences [23]. Those data variabilities and biases pose significant computational challenges ranging from the need for accurately detecting the nucleosome positioning to developing normalization methods to eliminate systematic biases.

Enormous efforts at the computational aspect have been put into processing the sequencing data and mapping the nucleosome organization and chromatin structure. These computational efforts include the development of novel computational methods, the application of sophisticated algorithms, the modeling of the distinct distribution of the sequencing data and the extraction of underlying data features. For example, NPS/iNPS [24,25] applied a gaussian convolution algorithm to detect inflection points to find candidate nucleosomes, and a novel iterative correction (ICE) algorithm [26] was developed for balancing the biases in Hi-C data. The computational algorithms are often further implemented as software tools for user-friendly and interactive interfaces or downloadable executable files in public code repositories, such as NucHMM [27] for identifying nucleosome states, JuiceBox [28] for visualizing and analyzing Hi-C data, and HiSIF [29] for detecting significant interacting fragments. In addition, a series of methods and tools are sometimes combined into a workflow/pipeline to achieve its meaningful biological output. For instance, HiC-Pro [30], a popular pipeline in Hi-C analysis, combines Bowtie2 (a mapping tool) [31], ICE, ggplot2 [32] (a visualization package) and other customized scripts for Hi-C data processing.

This review mainly focuses on the computational methods available for mapping the higher-resolution nucleosome and higher-order chromatin architectures, and the summaries of the features of the methods. While a detailed discussion of the underlying algorithms is beyond the scope of our survey, we will first discuss the methods and tools that can detect the nucleosomes in the genome, then demonstrate the computational methods for identifying 3D chromatin domains and interactions. We further illustrate computational approaches for integrating multi-omics data with Hi-C data and the advance of single-cell (sc)Hi-C data analysis. Our survey provides a comprehensive and valuable resource for biomedical scientists interested in studying nucleosome organization and chromatin structures and computational scientists interested in improving upon them.

2. Computational methods in detecting nucleosomes and characterizing nucleosome organization

Accumulated evidence showed that nucleosomes and nucleosome organization played essential roles in transcriptional regulation by interplaying with pioneer factors (PFs), transcription factors (TFs), and ATP-dependent chromatin remodelers [33,34]. Nucleosome organization, which is characterized by nucleosome positioning, spacing and phasing, is defined as the following (Fig. 1): 1) nucleosome positioning describes the consistency of the nucleosomes' position in the population of cells, where nucleosome position is referred to the location of the histone-octamer relative to DNA; 2) nucleosome spacing describes the distance between the dyads of two adjacent nucleosomes; 3) nucleosome phasing describes the degree of periodicity of the nucleosome array, where nucleosome array is referred to a string of nucleosomes with no more than 350 bp spacing. We didn't list nucleosome occupancy as a nucleosome organization feature because it is included in measuring nucleosome positioning. Since MNase-seq is the most prevalent and adaptable experimental method to detect genome-wide nucleosomes, in this section, we were mainly focused on summarizing computational methods for processing MNase-seq data. Unlike typical ChIP-seq data of TFs or histone modifications, MNase-seq data tend to have a more noised background signal. Thus classic ChIP-seq peak callers, such as MACS2 [35] and SPP [36], are not suitable for locating genome-wide nucleosomes. Although the basic idea for detecting nucleosomes is still the identification of enriched regions (peaks), the algorithmic design's goal is instead to improve the signal-to-noise ratio and refine the peak's width to a specific size, 147 ± 30 bp. Several methods based on the above algorithmic design have been developed, including ChIPseqR [37], iNPS [24], NOrMAL [38], NPS [25], Nseq [39], NucDe [40], NucHMM [27], NucHunter [41], NucleoFinder [42], nucleR [43], PING1/2 [44], PUFFIN [45]. Further, some other methods have been focused on identifying the changes of nucleosome positioning among cell types, e.g., BINOch [46], DANPOS/DANPOS2 [47], DiNuP [48] Dimnp [49] and Nucleosome Dynamics [50]. In addition, some methods, such as NucPosSimulator [51] and cplate [52], used Monte Carlo simulations and Template-Based Bayesian to identify alternative or averaged nucleosome positions in cell populations. Few methods, e.g., NucHMM and Nucleosome Dynamic, have lately been developed to further infer nucleosome organization, including nucleosome positioning, spacing and phasing. It is worth mentioning several other methods, TemplateFilter [53], a highly-cited method for analyzing nucleosome positioning from the tiling microarray data, though outdated but usually used as the control in performance comparison, and DeNOPA [54] and NucleoATAC [11], designed based on ATAC-seq data. A detailed summary of these methods is listed in Table 1. Concerning the comparisons for the performance of nucleosome

Table 1
Methods for detecting nucleosomes and characterizing nucleosome organization.

Tools	Description	Programming Language	Input data Type	Differential analysis	Quantitative Nucleosome organization Information	Webpage
BINoch	Identifying differential nucleosome occupancy regions with nucleosome stabilization–destabilization (NSD) score.	Python	Single-end (SE) Pair-end. (PE) MNase-seq	Yes	None	https://liulab-dfci.github.io/software/
ChIPseqR	A binding-event description model to locate nucleosomes, which is also flexible to handle other types of experiments.	R	SE MNase-seq ChIP-seq	No	Nucleosome positioning score (binding score in the package)	https://www.bioconductor.org/packages/release/bioc/html/ChIPseqR.html
DANPOS2	A toolkit for Dynamic Analysis of Nucleosome and Protein Occupancy by Sequencing, version 2.	Python	SE/PE MNase-seq ChIP-seq	Yes	Nucleosome positioning score	https://sites.google.com/site/danposdoc/
deNOPA	Decoding nucleosome positions with ATAC-seq data.	Python	ATAC-seq scATAC-seq	No	None	https://gitee.com/bxxu/denopa.
Dimnp	Identifying regions with differential nucleosome occupancy in multiple samples using Chi-squared test.	Python Matlab	Not Clear	Yes	None	https://bioinfo.seu.edu.cn/Nu_dynamics_data_public/
DiNuP	A systematic approach to identify regions of differential nucleosome positioning (RDNP).	Python	SE MNase-seq	Yes	Nucleosome positioning score	https://zhanglab.tongji.edu.cn/softwares/DiNuP/download.html
iNPS	An improved version of NPS, which outperforms latter one and provides additional nucleosome features.	Python	SE/PE MNase-seq	No	None	https://www.picb.ac.cn/hanlab/iNPS.html
NOrMAL	Using a modified Gaussian mixture model to identify nucleosome positions.	C++	SE/PE MNase-seq	No	Nucleosome positioning (Fuzziness)	https://github.com/antonpolishko/NOrMAL
NPS	A signal processing-based algorithm for identifying positioned nucleosomes from sequencing experiments at the nucleosome level.	Python	SE MNase-seq	No	None	https://liulab-dfci.github.io/software/
Nseq	A multithreaded Java application for finding positioned nucleosomes from sequencing data.	Java	SE MNase-seq	No	Nucleosome positioning score	https://github.com/songlab/Nseq
NucDe	A Non-homogeneous hidden-State model on first order differences for automatic detection of nucleosome positions.	R	SE MNase-seq MNase-Chip	No	None	https://pages.stat.wisc.edu/~keles/Software/demo_Nucde.pdf
NucDyn	Based on nucleR and aimed at comparing the reads of two MNase-seq experiments for nucleosome positioning and detecting significant inclusions, evictions and shifts.	R	SE/PE MNase-seq	Yes	Nucleosome positioning scorenucleosome phasing (Termed Periocity in paper) for specific gene (from TSS to TTS)	https://github.com/nucleosome-dynamics/NucDyn
NucHMM	A quantitative method modeling of nucleosome organization identifying functional nucleosome states, the nucleosome position is identified based on iNPS.	Python/C++	SE/PE MNase-seq ChIP-seq	No	Nucleosome positioning score Nucleosome spacing Nucleosome phasing	https://github.com/KunFang93/NucHMM
NucHunter	Predicting nucleosome positions with histone marks annotation from ChIP data.	Java	SE/PE ChIP-seq	No	Nucleosome positioning score (Fuzziness score)	https://epigen.molgen.mpg.de/nuchunter/
NucleoATAC	NucleoATAC is a python package for calling nucleosome positions and occupancy using ATAC-Seq data.	Python	PE ATAC-seq MNase-seq	No	Nucleosome positioning score	https://nucleoatac.readthedocs.io/en/latest/
NucleoFinder	A statistical approach for the detection of nucleosome positions. The authors claim it has fewer false positive detection than NPS and TemplateFilter.	R	SE MNase-seq MNase-Chip	No	Nucleosome spacing	https://sites.google.com/site/beckerjeremie/home/nucleofinder
nucleR	Using Non-parametric methods to detect nucleosome position. Its features include in situ visualization and exporting results to common genome browser formats.	R	SE/PE MNase-seq Tiling arrray data	No	Nucleosome positioning score	https://mmb.pcb.ub.es/nucleR/
NucTools	NucTools accounts for the continuous distribution of nucleosome occupancy.	Perl/Matlab	SE/PE MNase-	Yes	Nucleosome positioning scoreNucleosome	https://generegulation.org/nuctools/

Table 1 (*continued*)

Tools	Description	Programming Language	Input data Type	Differential analysis	Quantitative Nucleosome organization Information	Webpage
			seq ChIP-seq		spacing (NRL)	
NUCwave	A wavelet-based tool that is designed to evaluate the nucleosome-related experimental methods.	Python	SE/PE MNase-seq ChIP-seq Chemical mapping	No	None	https://nucleosome.usal.es/nucwave/
PING2	A Probabilistic inference method to identify nucleosome positioning.	R	SE/PE MNase-seq ChIP-seq	No	None	https://www.bioconductor.org/packages/release/bioc/html/PING.html
PuFFIN	A parameter-free method to construct genome-wide nucleosome maps from paired-end sequencing data.	Python	PE MNase-seq	No	Nucleosome positioning (Fuzziness)	https://github.com/ucrbioinfo/PuFFIN
TemplateFilter	Source code and executable files based pipeline for nucleosome positioning data processing.	Perl	SE MNase-seq	No	None	https://compbio.cs.huji.ac.il/NucPosition/TemplateFiltering/Home.html

peak calling tools and nucleosome characterization tools, the interested reader are referred to the review [55] and tools' original articles [24,27,44,45].

3. Computational methods in mapping 3D chromatin domains and interactions

Identification of chromatin compartments. Mammalian genomes are usually composed of two types of chromatin compartments, A and B compartments [17], corresponding to open and closed chromatin, respectively. Compartments were first derived and defined by the sign of the first principal component analysis (PCA) on transformed Hi-C matrices. Many tools, such as JuiceBox [56], HOMER [57] and Fan-C [58], implemented original compartment analysis in their suite. It was further extended to a finer sub-compartment on Hi-C data with a very high sequencing coverage such as Hi-C GM12878 data [59] by SNIPER developed by Xiong and Ma [60], based on a denoising autoencoder and a multilayer perceptron classifier. A recent new tool, Calder [61] was able to identify multi-scale sub-compartments at variable data resolutions, even in experiments with relatively low resolution.

Prediction of TADs. TAD is one of the key 3D-genome structures that control gene regulation [62]. So far, more than two dozen [59,63–66] programs or algorithms have been developed to predict TADs. Most of TAD prediction methods can be grouped into four major categories: one-dimension linear score method (e.g., TopDom [67]), two-dimension clustering method (e.g., MSTD [68]), feature method (e.g., pTADS [69]), and statistical method (e.g., HiCseg [70]). Recently, a systematic comparison of 22 computational methods for predicting TADs was carried out [66], providing a detailed description illustrating the pros and cons of each of the 22 methods. This extensive evaluation of TAD prediction tools is an excellent resource that safe-guides users in selecting a suitable tool or method for the TAD prediction, based on the desired experimental design and biological question. Although TADs are commonly conserved at a larger scale among different conditions, Sub-TADs within each large TAD are believed to be dynamically changed between the conditions [62]. Some of TAD prediction tools are specially designed to detect both the TAD and the sub-TAD, including GMAP [71], HiTAD [72], Arrowhead [73] and TADtree [74]. A more recent tool, OnTAD [75], implemented a two-level approach to generate a hierarchical TAD organization that best fits the input Hi-C interaction matrix. It is reported to performs better than other tools and reveals several interesting biological phenomena related to a TAD hierarchy.. It is worth mentioning that a set of methods were particularly developed for some *variant Hi-C techniques such as* ChIA-PET [76], capture-Hi-C [77]), HiChIP [78] and Split-Pool Recognition of Interactions by Tag Extension (SPRITE) [79] due to their sparse or asymmetric interaction matrices. GRiNCH [80] is one such tool based on a non-negative matrix factorization and graph regularization, to identify TADs of genome organization from sparse chromosome interaction matrices.

Prediction of differential TAD/TAD boundaries. Besides the computational prediction of TADs, there is also a need for computational methods to investigate dynamical changes of TADs between the conditions. For example, researchers have recently observed differential responses at TADs, looping genes and expressed genes in endocrine-resistant breast cancer cells [81]. A new algorithmic approach DADo [82], has been proposed to identify such differentially active domains between the two conditions, providing complementary information to a general differential expression analysis. However, there are not many publicly available tools for the aforementioned task. In the future, more tools are needed to perform the differential analysis of chromatin domains between different conditions.

Identification of chromatin interactions and loops. In addition to mapping 3D chromatin domains, computational methods for identifying chromatin interactions or loops have also been developed quickly in the past few years. A probabilistic mixture model [83] was proposed to identify promoter-enhancer interactions from Hi-C data about five years ago. However, the sequencing depth required to obtain high-resolution interactions from Hi-C data is too big. A new tool, HIFI [84], based on a density estimation algorithm was thus designed to detect high resolution (at the restriction-fragment scale) chromatin interactions from Hi-C data. Alternately, a combination of low input "easy Hi-C" protocol for 3D genome mapping and a new analysis pipeline (HiCorr) [85] for Hi-C bias-correction at high resolution, were able to detect enhancer-promoter loops at sub-TAD level. Besides, FitHiC1/2 [86,87] implemented the statistical confidence estimation method to detect loops. And a novel computational and statistical method (HiSIF) [29] was proposed to identify genome-wide chromatin loops in Hi-C data with high resolution, by using a two-tier module (quality

control and classification). HiSIF detected genes with enhanced loops showing worse survival in endocrine-treated breast cancer patients. FitHiChIP [88] is a computational tool for calling chromatin loops from Hi-C/HiChiP/PLAC-seq data with a better performance. More recently, MUSTACHE [89] employed scale-space theory in computer vision to detect loops in Hi-C contact matrix. For readers who are interested in studying the algorithm and comparing the performance of the Hi-C related tools, recent review papers [66,90,91] are referred.

4. Computational approaches for integrating multi-omics data including Hi-C data

The availability of Hi-C data and other genome-wide omics data from 4D Nucleome [92], ENCODE [93] and other resources provide an excellent opportunity to develop computational approaches to integrating multi-omics data from the same or different sources. Some pioneering studies explored integrative computational methods in the context of 3D genome organization. Lan etc. [94] used a Mixture Poisson Regression Model and a power-law decay background to identify a highly specific set of interacting genomic regions from publicly available K562 Hi-C data, sequentially applied hierarchical clustering to ChIP-seq data of nine histone marks and an Apriori algorithm to ChIP-seq of 45 transcription factors, and collectively classified 12 different clusters of interacting loci categorized into two types of chromatin linkages. Whalen et al. [95] developed an ensemble machine learning pipeline, TargetFinder, for predicting enhancer-promoter interactions by integrating Hi-C interactions with epigenetic annotations from Segway and ChromHMM. Pancaldi et al. [96], inspired by a network topology metric used in social sciences, applied the metric to integrate epigenomic data and high-resolution promoter capture Hi-C and Hi-Cap data as well as ChIA-PET data to investigate promoter-centered chromatin interaction networks. Wang et al. [97] proposed a computational method called SPIN which integrated TSA-seq, DamID and Hi-C data from K562 cells in a unified framework based on a hidden Markov random field. They revealed a genome-wide intra-nuclear chromosome positioning and nuclear compartmentalization relative to multiple nuclear structures, such as nuclear lamina, speckles and nucleoli. Recently, a new method pTADs [69] was developed to identify TAD boundary and strength, by integrating both DNA sequence-based features (e.g., DNA shape and TF binding motif occurrence) and epigenetic profile information (e.g., CTCF, H3K36me3 and H3K20me1).

5. Advances in scHi-C computational analyses

Although tremendous progress in reconstructing the 3D chromatin structure based on population-averaged Hi-C data [98], single-cell Hi-C (scHi-C) protocols have newly been developed to identify 3D chromatin architecture at single-cell resolution [99–102], in which it has capability to delineate the 3D-regulated heterogeneity in population cells. For instance, the organization of zygote chromatin [103], the nuclear changes of stem cell differentiation [104], and single-allele chromatin interactions [105,106] have been fully examined by scHi-C technique. A crucial issue with scHi-C analysis is the inherent sparsity of the contact matrices and the technical noise mainly due to low amounts of starting material often resulting in variable capture efficiencies. Quality control of sequencing data is crucial to avoid technical artifacts. Despite of these challenges, new sets of computational methods have been developed for processing scHi-C data to reconstruct single-cell 3D chromatin structures [107–109], to impute the chromosome contact matrices [110–112], to identify TAD-like domains [113], to classify single cells [114], to identify chromatin loops [115], and to provide toolbox of scHi-C [116].

6. Conclusions and future perspectives

In summary, our survey provides a comprehensive and valuable resource for biomedical scientists interested in studying nucleosome organization and chromatin structures, as well as for computational scientists who are interested in improving upon them. Despite such abundant resources, there is still an opportunity to develop more tools for integrating multi-source information and multi-level approaches in analyzing 3D transcriptional regulation. We also urgently need computational tools which could directly interpret the functionality of the nucleosome organization and 3D chromatin structure in various diseases. Another promising area worthy of attention is to incorporate molecular imaging data such as 3D super-resolution imaging (3D-SIM) [117] and genome architecture mapping (GAM) [118] to better derive the nucleosome and chromatin structures.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partly supported by National Institutes of Health (NIH R01GM114142 and U54CA217297 to V.X.J.) and NDSU Startup Award to L.L. J.W. was supported by South-Eastern Norway Regional Health Authority [HSØ 2018107] and the Norwegian Research Council NOTUR project (nn4605k).

References

[1] Luger K et al. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 1997;389(6648):251–60.
[2] Splinter E et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev 2006;20(17):2349–54.
[3] Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell 2009;137(7):1194–211.
[4] Weintraub AS et al. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell 2017;171(7):1573–1588 e28.
[5] Miele A, Dekker J. Long-range chromosomal interactions and gene regulation. Mol Biosyst 2008;4(11):1046–57.
[6] Dekker J. Gene regulation in the third dimension. Science 2008;319(5871):1793–4.
[7] Zheng H, Xie W. The role of 3D genome organization in development and cell differentiation. Nat Rev Mol Cell Biol 2019;20(9):535–50.
[8] Baldi S et al. Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. Nat Struct Mol Biol 2018;25(9):894–901.
[9] Kilgore JA et al. Single-molecule and population probing of chromatin structure using DNA methyltransferases. Methods 2007;41(3):320–32.
[10] Brogaard K et al. A map of nucleosome positions in yeast at base-pair resolution. Nature 2012;486(7404):496–501.
[11] Schep AN et al. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome Res 2015;25(11):1757–70.
[12] Ishii H, Kadonaga JT, Ren B. MPE-seq, a new method for the genome-wide analysis of chromatin structure. Proc Natl Acad Sci U S A 2015;112(27):E3457–65.
[13] Dekker J et al. Capturing chromosome conformation. Science 2002;295(5558):1306–11.
[14] Simonis M et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet 2006;38(11):1348–54.
[15] Dostie J et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res 2006;16(10):1299–309.
[16] Hughes JR et al. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat Genet 2014;46(2):205–12.

[17] Lieberman-Aiden E et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009;326 (5950):289–93.
[18] Kalhor R et al. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat Biotechnol 2011;30(1):90–8.
[19] Hsieh TH et al. Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell 2015;162(1):108–19.
[20] Horz W, Altenburger W. Sequence specific cleavage of DNA by micrococcal nuclease. Nucleic Acids Res 1981;9(12):2643–58.
[21] Chereji RV, Bryson TD, Henikoff S. Quantitative MNase-seq accurately maps nucleosome occupancy levels. Genome Biol 2019;20(1):198.
[22] Gutierrez G et al. Subtracting the sequence bias from partially digested MNase-seq data reveals a general contribution of TFIIS to nucleosome positioning. Epigenet Chromat 2017;10(1):58.
[23] Yaffe E, Tanay A. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. Nat Genet 2011;43(11):1059–65.
[24] Chen W et al. Improved nucleosome-positioning algorithm iNPS for accurate nucleosome positioning from sequencing data. Nat Commun 2014;5:4909.
[25] Zhang Y et al. Identifying positioned nucleosomes with epigenetic marks in human from ChIP-Seq. BMC Genomics 2008;9:537.
[26] Imakaev M et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 2012;9(10):999–1003.
[27] Fang K et al. NucHMM: a method for quantitative modeling of nucleosome organization identifying functional nucleosome states distinctly associated with splicing potentiality. Genome Biol 2021;22(1):250.
[28] Robinson JT et al. Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. Cell Syst 2018;6(2):256–258 e1.
[29] Zhou Y et al. Modeling and analysis of Hi-C data by HiSIF identifies characteristic promoter-distal loops. Genome Med 2020;12(1):69.
[30] Servant N et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol 2015;16:259.
[31] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9(4):357–9.
[32] Villanueva, R.A.M. Z.J. Chen, *ggplot2: elegant graphics for data analysis*; 2019, Taylor & Francis.
[33] Schones DE et al. Dynamic regulation of nucleosome positioning in the human genome. Cell 2008;132(5):887–98.
[34] Struhl K, Segal E. Determinants of nucleosome positioning. Nat Struct Mol Biol 2013;20(3):267–73.
[35] Zhang Y et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9 (9):R137.
[36] Kharchenko PV, Tolstorukov MY, Park PJ. Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat Biotechnol 2008;26(12):1351–9.
[37] Humburg P et al. ChIPseqR: analysis of ChIP-seq experiments. BMC Bioinf 2011;12:39.
[38] Polishko A et al. NORMAL: accurate nucleosome positioning using a modified Gaussian mixture model. Bioinformatics 2012;28(12):i242–9.
[39] Nellore A et al. NSeq: a multithreaded Java application for finding positioned nucleosomes from sequencing data. Front Genet 2012;3:320.
[40] Kuan PF et al. A non-homogeneous hidden-state model on first order differences for automatic detection of nucleosome positions. Stat Appl Genet Mol Biol 2009;8. Article29.
[41] Mammana A, Vingron M, Chung HR. Inferring nucleosome positions with their histone mark annotation from ChIP data. Bioinformatics 2013;29 (20):2547–54.
[42] Becker J et al. NucleoFinder: a statistical approach for the detection of nucleosome positions. Bioinformatics 2013;29(6):711–6.
[43] Flores O, Orozco M. nucleR: a package for non-parametric nucleosome positioning. Bioinformatics 2011;27(15):2149–50.
[44] Woo S, Ping,, et al. 2.0: an R/Bioconductor package for nucleosome positioning using next-generation sequencing data. Bioinformatics 2013;29 (16):2049–50.
[45] Polishko A et al. PuFFIN–a parameter-free method to build nucleosome maps from paired-end reads. BMC Bioinf 2014;15(Suppl 9):S11.
[46] Meyer CA et al. BINOCh: binding inference from nucleosome occupancy changes. Bioinformatics 2011;27(13):1867–8.
[47] Chen K et al. DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res 2013;23(2):341–51.
[48] Fu K et al. DiNuP: a systematic approach to identify regions of differential nucleosome positioning. Bioinformatics 2012;28(15):1965–71.
[49] Liu L et al. An approach of identifying differential nucleosome regions in multiple samples. BMC Genomics 2017;18(1):135.
[50] Buitrago D et al. Nucleosome Dynamics: a new tool for the dynamic analysis of nucleosome positioning. Nucleic Acids Res 2019;47(18):9511–23.
[51] Schopflin R et al. Modeling nucleosome position distributions from experimental nucleosome positioning maps. Bioinformatics 2013;29 (19):2380–6.
[52] Zhou X et al. A computational approach to map nucleosome positions and alternative chromatin states with base pair resolution. Elife 2016;5.
[53] Zhang X et al. Probabilistic inference for nucleosome positioning with MNase-based or sonicated short-read data. PLoS ONE 2012;7(2):e32095.
[54] Xu B et al. DeNOPA: decoding nucleosome positions sensitively with sparse ATAC-seq data. Brief Bioinform 2021.
[55] Quintales L, Vazquez E, Antequera F. Comparative analysis of methods for genome-wide nucleosome cartography. Brief Bioinform 2015;16(4):576–87.
[56] Durand NC et al. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst 2016;3(1):99–101.
[57] Heinz S et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38(4):576–89.
[58] Kruse K, Hug CB, Vaquerizas JM. FAN-C: a feature-rich framework for the analysis and visualisation of chromosome conformation capture data. Genome Biol 2020;21(1):303.
[59] Rao SS et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159(7):1665–80.
[60] Xiong K, Ma J. Revealing Hi-C subcompartments by imputing inter-chromosomal chromatin interactions. Nat Commun 2019;10(1):5069.
[61] Liu Y et al. Systematic inference and comparison of multi-scale chromatin sub-compartments connects spatial organization to cell phenotypes. Nat Commun 2021;12(1):2439.
[62] Dixon JR, Gorkin DU, Ren B. Chromatin Domains: The Unit of Chromosome Organization. Mol Cell 2016;62(5):668–80.
[63] Dixon JR et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 2012;485(7398):376–80.
[64] Sexton T et al. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 2012;148(3):458–72.
[65] Filippova D et al. Identification of alternative topological domains in chromatin. Algorithms Mol Biol 2014;9:14.
[66] Zufferey M et al. Comparison of computational methods for the identification of topologically associating domains. Genome Biol 2018;19(1):217.
[67] Shin H et al. TopDom: an efficient and deterministic method for identifying topological domains in genomes. Nucleic Acids Res 2016;44(7):e70.
[68] Ye Y, Gao L, Zhang S. MSTD: an efficient method for detecting multi-scale topological domains from symmetric and asymmetric 3D genomic maps. Nucleic Acids Res 2019;47(11):e65.
[69] Wang Y et al. TAD boundary and strength prediction by integrating sequence and epigenetic profile information. Brief Bioinform 2021;22(5).
[70] Levy-Leduc C et al. Two-dimensional segmentation for analyzing Hi-C data. Bioinformatics 2014;30(17):i386–92.
[71] Yu W, He B, Tan K. Identifying topologically associating domains and subdomains by Gaussian Mixture model And Proportion test. Nat Commun 2017;8(1):535.
[72] Wang XT, Cui W, Peng C. HiTAD: detecting the structural and functional hierarchies of topologically associating domains from chromatin interactions. Nucleic Acids Res 2017;45(19):e163.
[73] Durand NC et al. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst 2016;3(1):95–8.
[74] Weinreb C, Raphael BJ. Identification of hierarchical chromatin domains. Bioinformatics 2016;32(11):1601–9.
[75] An L et al. OnTAD: hierarchical domain structure reveals the divergence of activity among TADs and boundaries. Genome Biol 2019;20(1):282.
[76] Zhang Y et al. Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. Nature 2013;504(7479):306–10.
[77] Javierre BM et al. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell 2016;167 (5):1369–1384 e19.
[78] Mumbach MR et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods 2016;13(11):919–22.
[79] Quinodoz SA et al. Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. Cell 2018;174(3):744–757 e24.
[80] Lee DI, Roy S. GRiNCH: simultaneous smoothing and detection of topological units of genome organization from sparse chromatin contact count matrices with matrix factorization. Genome Biol 2021;22(1):164.
[81] Yang Y et al. The 3D genomic landscape of differential response to EGFR/HER2 inhibition in endocrine-resistant breast cancer cells. Biochim Biophys Acta Gene Regul Mech 2020;1863(11):194631.
[82] Zufferey M et al. Systematic assessment of gene co-regulation within chromatin domains determines differentially active domains across human cancers. Genome Biol 2021;22(1):218.
[83] Ron G et al. Promoter-enhancer interactions identified from Hi-C data using probabilistic models and hierarchical topological domains. Nat Commun 2017;8(1):2237.
[84] Cameron CJ, Dostie J, Blanchette M. HIFI: estimating DNA-DNA interaction frequency from Hi-C data at restriction-fragment resolution. Genome Biol 2020;21(1):11.
[85] Lu L et al. Robust Hi-C Maps of Enhancer-Promoter Interactions Reveal the Function of Non-coding Genome in Neural Development and Diseases. Mol Cell 2020;79(3):521–534 e15.
[86] Ay F, Bailey TL, Noble WS. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. Genome Res 2014;24(6):999–1011.
[87] Kaul A, Bhattacharyya S, Ay F. Identifying statistically significant chromatin contacts from Hi-C data with FitHiC2. Nat Protoc 2020;15(3):991–1012.
[88] Bhattacharyya S et al. Identification of significant chromatin contacts from HiChIP data by FitHiChIP. Nat Commun 2019;10(1):4221.
[89] Roayaei Ardakany A et al. Mustache: multi-scale detection of chromatin loops from Hi-C and Micro-C maps using scale-space representation. Genome Biol 2020;21(1):256.
[90] Forcato M et al. Comparison of computational methods for Hi-C data analysis. Nat Methods 2017;14(7):679–85.

[91] Lyu H, Liu E, Wu Z. Comparison of normalization methods for Hi-C data. Biotechniques 2020;68(2):56–64.
[92] Dekker J et al. The 4D nucleome project. Nature 2017;549(7671):219–26.
[93] Ecker JR et al. Genomics: ENCODE explained. Nature 2012;489(7414):52–5.
[94] Lan X et al. Integration of Hi-C and ChIP-seq data reveals distinct types of chromatin linkages. Nucleic Acids Res 2012;40(16):7690–704.
[95] Whalen S, Truty RM, Pollard KS. Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin. Nat Genet 2016;48 (5):488–96.
[96] Pancaldi V et al. Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity. Genome Biol 2016;17(1):152.
[97] Wang Y et al. SPIN reveals genome-wide landscape of nuclear compartmentalization. Genome Biol 2021;22(1):36.
[98] Imakaev, M.V., Fudenberg, G. Mirny, L.A. *Modeling chromosomes: Beyond pretty pictures.* FEBS Lett; 2015. **589** (20 Pt A): p. 3031-6.
[99] Nagano T et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 2013;502(7469):59–64.
[100] Ramani V et al. Massively multiplex single-cell Hi-C. Nat Methods 2017;14 (3):263–6.
[101] Stevens TJ et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature 2017;544(7648):59–64.
[102] Nagano T et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. Nature 2017;547(7661):61–7.
[103] Flyamer IM et al. Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature 2017;544 (7648):110–4.
[104] Bonora G et al. Single-cell landscape of nuclear configuration and gene expression during stem cell differentiation and X inactivation. Genome Biol 2021;22(1):279.
[105] Allahyar A et al. Enhancer hubs and loop collisions identified from single-allele topologies. Nat Genet 2018;50(8):1151–60.
[106] Oudelaar AM et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. Nat Genet 2018;50 (12):1744–51.
[107] Rosenthal M et al. Bayesian Estimation of Three-Dimensional Chromosomal Structure from Single-Cell Hi-C Data. J Comput Biol 2019;26(11):1191–202.
[108] Zhu H, Wang Z. SCL: a lattice-based approach to infer 3D chromosome structures from single-cell Hi-C data. Bioinformatics 2019;35(20):3981–8.
[109] Meng L et al. Si-C is a method for inferring super-resolution intact genome structure from single-cell Hi-C data. Nat Commun 2021;12(1):4369.
[110] Liu J et al. Unsupervised embedding of single-cell Hi-C data. Bioinformatics 2018;34(13):i96–i104.
[111] Zhou J et al. Robust single-cell Hi-C clustering by convolution- and random-walk-based imputation. Proc Natl Acad Sci U S A 2019;116(28):14011–8.
[112] Zhang R, Zhou T, Ma J. Multiscale and integrative single-cell Hi-C analysis with Higashi. Nat Biotechnol 2022;40(2):254–61.
[113] Li X et al. DeTOKI identifies and characterizes the dynamics of chromatin TAD-like domains in a single cell. Genome Biol 2021;22(1):217.
[114] Wu H et al. scHiCStackL: a stacking ensemble learning-based method for single-cell Hi-C classification using cell embedding. Brief Bioinform 2022;23 (1).
[115] Yu M et al. SnapHiC: a computational pipeline to identify chromatin loops from single-cell Hi-C data. Nat Methods 2021;18(9):1056–9.
[116] Li X et al. scHiCTools: A computational toolbox for analyzing single-cell Hi-C data. PLoS Comput Biol 2021;17(5):e1008978.
[117] Chowdhury R et al. The use of molecular imaging combined with genomic techniques to understand the heterogeneity in cancer metastasis. Br J Radiol 2014;87(1037):20140065.
[118] Winick-Ng W et al. Cell-type specialization is encoded by specific chromatin topologies. Nature 2021;599(7886):684–91.